

Unsupervised Scene Sketch to Photo Synthesis

Jiayun Wang, Sangryul Jeon, Stella X. Yu,
UC Berkeley/ICSI

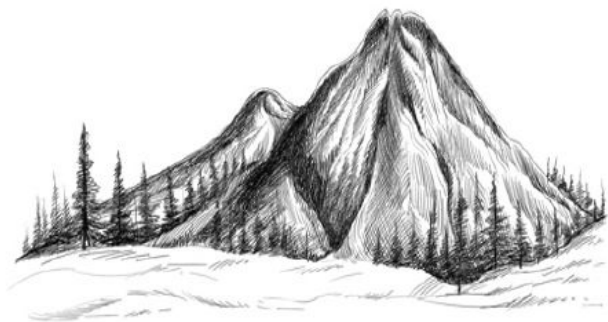


Xi Zhang, Himanshu Arora, Yu Lou
Amazon



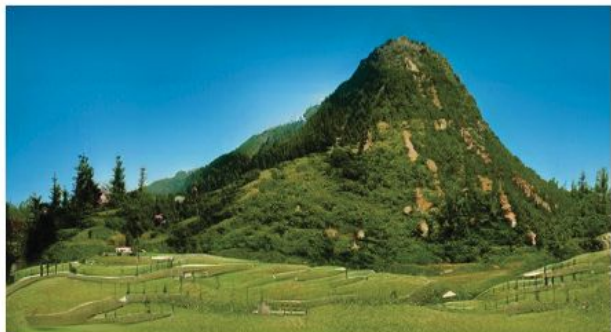
Sketch-Based Scene Generation

sketch



Sketch-Based Scene Generation

synthesized photo



Sketch-Based Scene Generation

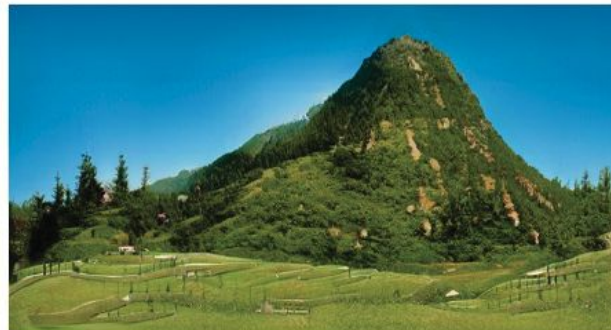
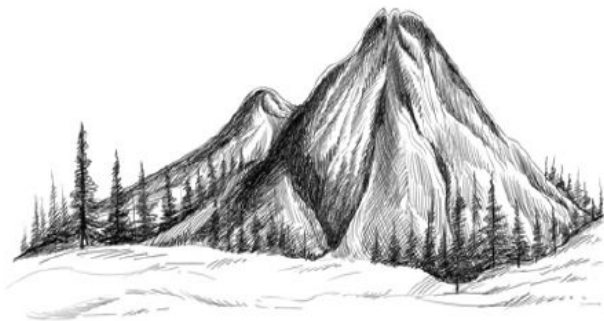
sketch



style reference

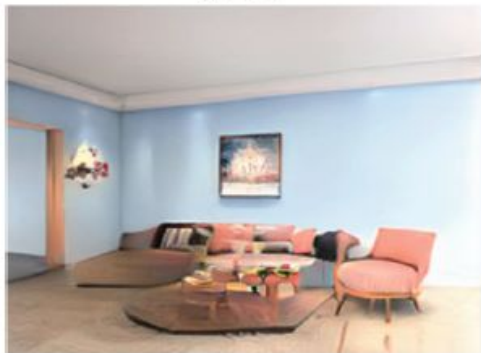


synthesized photo



Sketch-Based Scene Edit

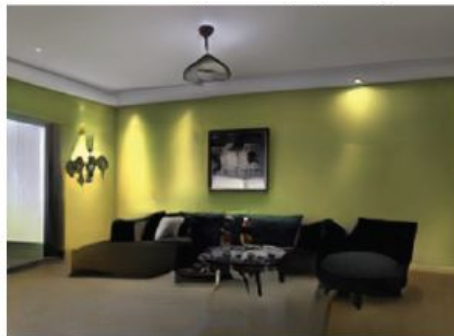
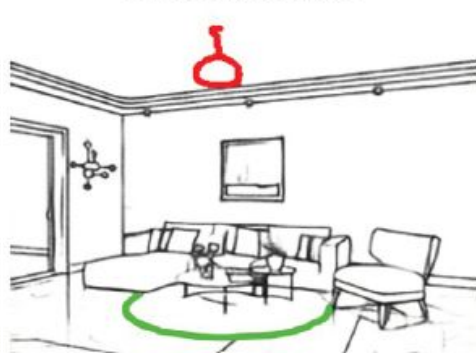
photo



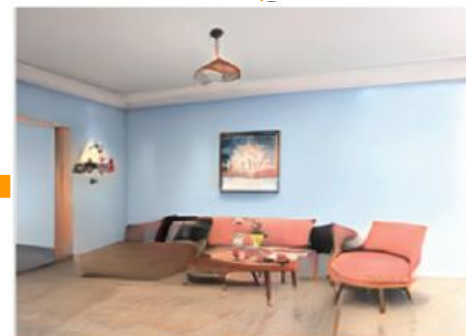
synthesized sketch



modified sketch

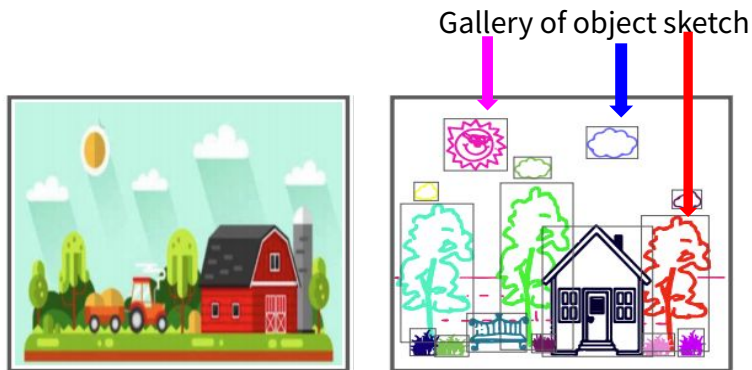


new style

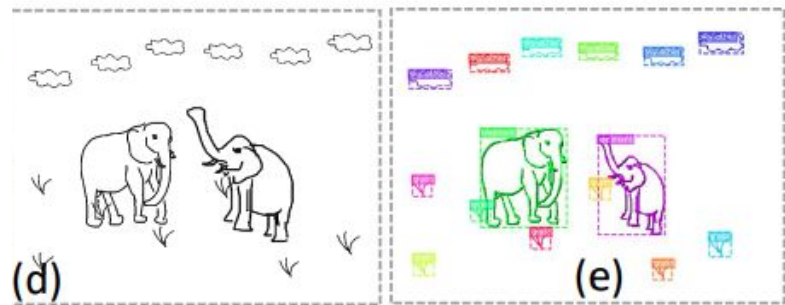


edited photo

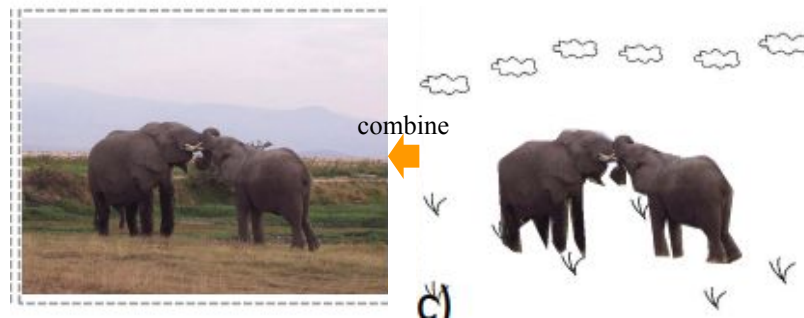
Existing Work: Break to Objects, Synthesize, then Combine



SkethyScene, ECCV 18

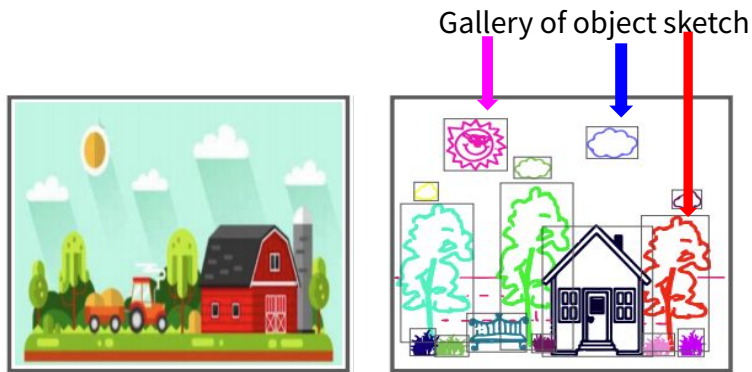


object sketch to photo



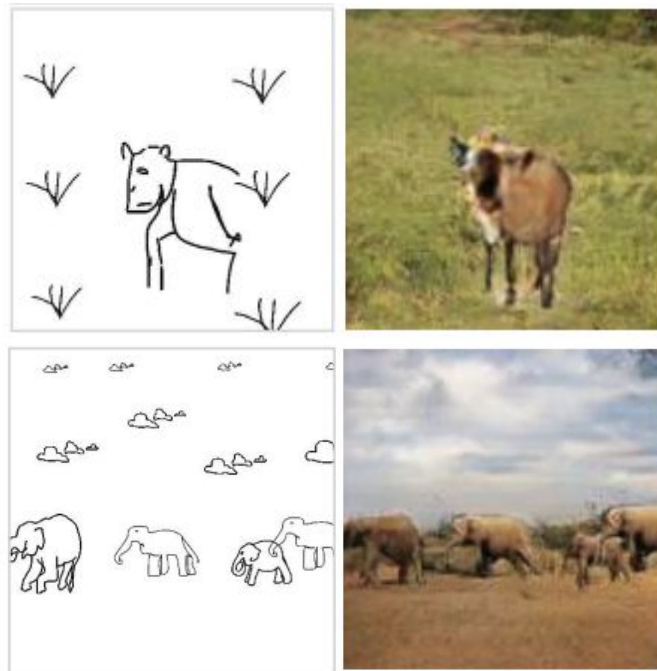
SkethyCOCO, CVPR 20

Existing Work: Break to Objects, Synthesize, then Combine



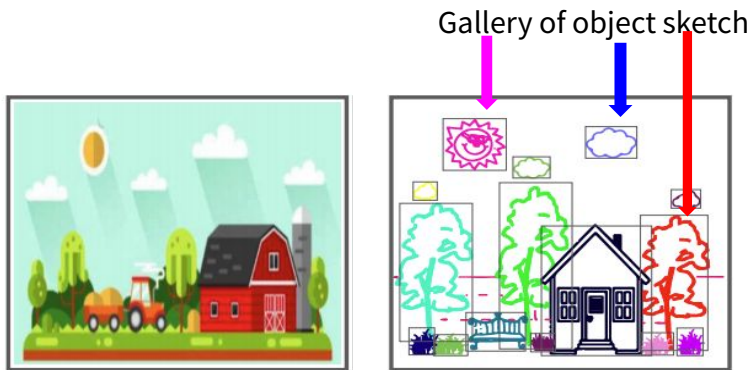
SkethyScene, 2018

Artifacts come from combination...
Can we directly learn from images?



SkethyCOCO, 2020

Learning Directly from Photos



SkethyScene, 2018



ours

Takeaway: the holistic scene structures need to be preserved rather than being broken into pieces

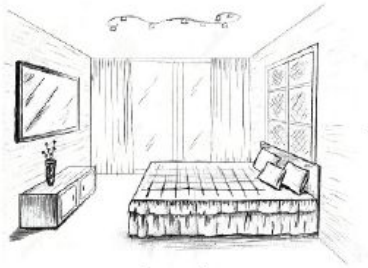
Learning Directly from Photos

standardization

Training

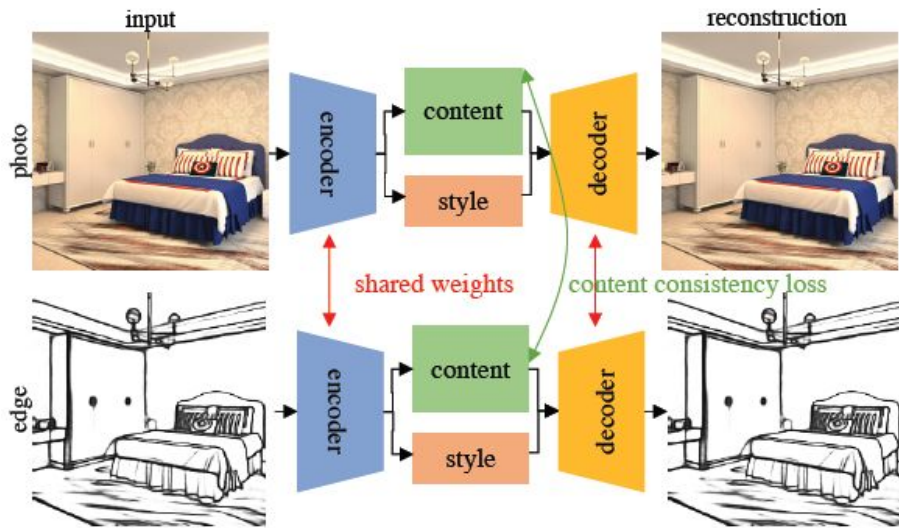


photo
Inference

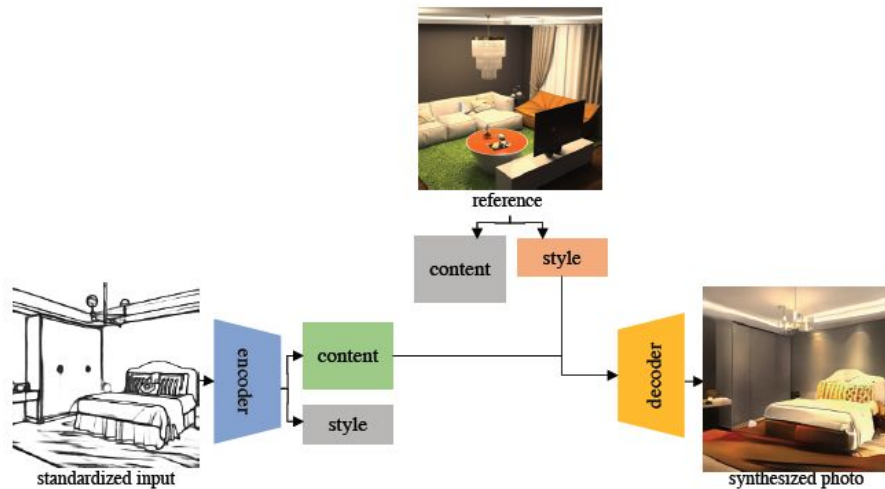


sketch

Training Pipeline



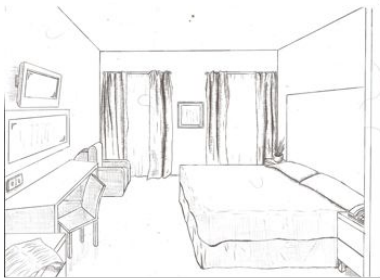
Step1: auto-encoding



Step2: fine-tuning with unpaired data

Results - Room Sketches

style
sketch



Results - Church Sketches

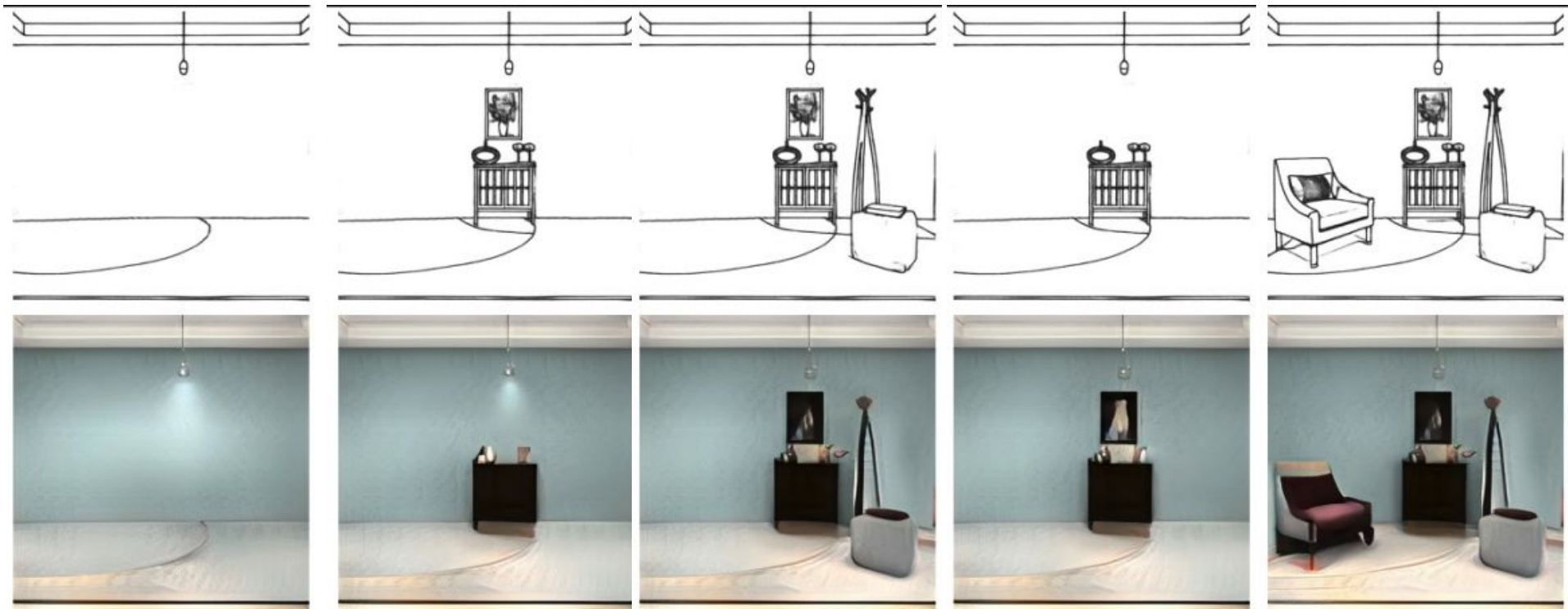
style
sketch



Results - Style Interpolation

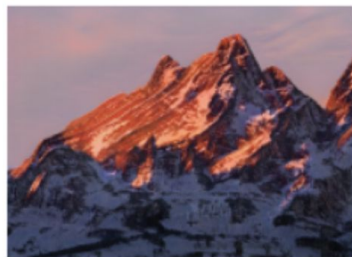


Results - Scene Editing



Perceptual Study: Can Generated Photo Fool People?

Which one is a real photo?



A



B

(%)	indoor scene	church	mountain	mean
ours	25.00	44.3	48.9	39.4
SAE [28]	10.0	6.6	20.0	12.2

51 subjects, 2k judgements

Questions on realism, style and content matching